

OSCE no Brasil: trajetória, fundamentos e perspectivas para a avaliação das competências clínicas

OSCE in Brazil: history, foundations, and perspectives for the assessment of clinical competencies

Marcos Carvalho Borges¹  marcosborges@fmrp.usp.br

Valdes Roberto Bollela¹  vbollela@fmrp.usp.br

Luiz Ernesto de Almeida Troncon¹  ledatron@fmrp.usp.br

RESUMO

Introdução: A educação médica baseada em competências exige métodos capazes de avaliar, de forma válida e confiável, o desempenho clínico do estudante. O *Objective Structured Clinical Examination* (OSCE) consolidou-se internacionalmente como um dos principais instrumentos para aferir o nível *shows how* (demonstrar como faz) da pirâmide de Miller.

Desenvolvimento: Este ensaio discute o lugar do OSCE na avaliação das competências clínicas, revisa sua trajetória internacional e latino-americana, resgata o protagonismo brasileiro na introdução do método na América Latina e analisa seus usos em avaliações formativas, somativas, informativas (diagnósticas) e de proficiência. Sustenta-se que o debate brasileiro atual não deve contrapor a necessidade de avaliação prática à inviabilidade de um OSCE de âmbito nacional centralizado. Alternativamente, o elevadíssimo número de escolas médicas no Brasil, com forte heterogeneidade institucional, faz com que um modelo híbrido pareça mais factível: avaliação cognitiva de âmbito nacional associada à avaliação descentralizada das habilidades clínicas nas próprias escolas ou em escolas organizadas em microrregiões, mas sob matriz nacional, documentação padronizada e estrita regulação e auditoria externas. Argumenta-se que a experiência histórica do OSCE no Brasil, especialmente a partir das iniciativas pioneiras da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo, oferece evidências valiosas sobre potencialidades e limites do método, incluindo custo operacional, necessidade de capacitação docente, treinamento de pacientes padronizados, *blueprinting*, definição de padrões de desempenho e análise psicométrica.

Conclusão: Defende-se que o futuro da avaliação prática no Brasil dependa menos da simples expansão quantitativa do OSCE e mais da construção de uma governança nacional de qualidade, baseada em bancos compartilhados de estações, consórcios interinstitucionais, revisão pós-implementação e integração do exame a sistemas de avaliação programática.

Palavras-chave: Avaliação Educacional; Avaliação de Desempenho Profissional; Competência Clínica; Estudantes de Medicina.

ABSTRACT

Introduction: Competency-based medical education requires methods capable of assessing students' clinical performance in a valid and reliable manner. The *Objective Structured Clinical Examination* (OSCE) has become internationally established as one of the main instruments for assessing the *shows how* level of Miller's pyramid.

Development: This essay discusses the role of the OSCE in the assessment of clinical competencies, reviews its international and Latin American trajectory, highlights the pioneering Brazilian contribution to the introduction of the method in Latin America, and analyses its use in formative, summative, informative (diagnostic), and proficiency assessments. We argue that the current Brazilian debate should not frame the need for practical assessment as incompatible with the logistical impracticality of a single nationwide centralized OSCE. Given the extremely large number of medical schools in Brazil, together with marked institutional heterogeneity, a hybrid model appears more feasible: a national cognitive assessment combined with decentralized assessment of clinical skills conducted within individual medical schools or regional consortia, but under a national competency framework, standardized documentation requirements, and strict external regulation and auditing. We further argue that the historical experience of the OSCE in Brazil, particularly the pioneering initiatives developed at the Ribeirão Preto Medical School, University of São Paulo, provides valuable evidence regarding both the strengths and limitations of the method, including operational costs, the need for faculty development, standardized patient training, *blueprinting*, standard setting, and psychometric analysis.

Conclusion: We contend that the future of practical assessment in Brazil depends less on the mere quantitative expansion of the OSCE and more on the establishment of a national quality governance structure based on shared station banks, interinstitutional consortia, post-implementation review, and integration of the examination into programmatic assessment systems.

Keywords: Educational Measurement; Clinical Skill; Clinical Competence; Students.

¹ Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, São Paulo, Brasil.

Editora-chefe: Rosiane Viana Zuza Diniz.

Editor associado: Fernando Almeida.

Recebido em 20/05/26; Aceito em 23/06/26.

Avaliado pelo processo de double blind review.

INTRODUÇÃO

A avaliação do estudante de Medicina vem sendo progressivamente compreendida como um processo contínuo de obtenção e interpretação de informações sobre o desempenho, voltado a pelo menos três finalidades: fomentar a aprendizagem, sustentar decisões de progressão e certificação e produzir evidências sobre a qualidade da formação oferecida pelas escolas médicas¹⁻⁴. No contexto contemporâneo, essas finalidades não são cumpridas de modo satisfatório por um único instrumento, mas por sistemas de avaliação capazes de combinar múltiplas fontes de evidência e de integrar seus resultados em decisões educacionais defensáveis^{1, 5-7}.

A difusão da educação baseada em competências reforçou essa necessidade. Quando o objetivo da formação é desenvolver a capacidade de mobilizar conhecimentos, habilidades, atitudes e valores em situações profissionais concretas, a avaliação não pode restringir-se ao domínio cognitivo. Ela precisa captar como o estudante atua, comunica-se, examina pacientes, prioriza problemas, toma decisões e se comporta em situações clínicas variadas⁵⁻⁷. Nessa perspectiva, o exame clínico objetivo estruturado – ou *Objective Structured Clinical Examination* (OSCE) –, consolidou-se como método de referência para a avaliação padronizada de habilidades e competências clínicas em ambiente simulado, com importante papel em programas de avaliação e em processos decisórios de maior relevância⁷⁻¹⁴.

No Brasil, essa discussão ganhou renovada importância com o aprofundamento do debate sobre a qualidade da formação médica, a revisão das Diretrizes Curriculares Nacionais e a retomada de mecanismos externos de avaliação, entre os quais se destaca o Exame Nacional de Avaliação da Formação Médica (Enamed)¹⁵⁻¹⁹. Ao mesmo tempo, o país convive com um sistema formador amplo, heterogêneo e em rápida expansão, o que torna ainda mais desafiadora a tarefa de construir avaliações válidas, confiáveis e factíveis para todos os concluintes.

Neste ensaio, revisamos os fundamentos que justificam o uso do OSCE na avaliação das competências clínicas, resgatamos aspectos da sua introdução e difusão no Brasil e discutimos possibilidades de articulação entre avaliações locais de desempenho e mecanismos nacionais de regulação. Sustentamos que o OSCE deve ser universalizado na formação médica brasileira, mas não por meio de um exame prático nacional único e centralizado. Em um país de dimensões continentais e forte heterogeneidade institucional, parece mais realista combinar uma prova cognitiva de abrangência nacional com avaliação prática descentralizada, realizada

nas escolas médicas ou em arranjos regionais, sob matriz nacional, documentação padronizada, garantias de qualidade psicométrica e auditoria externa.

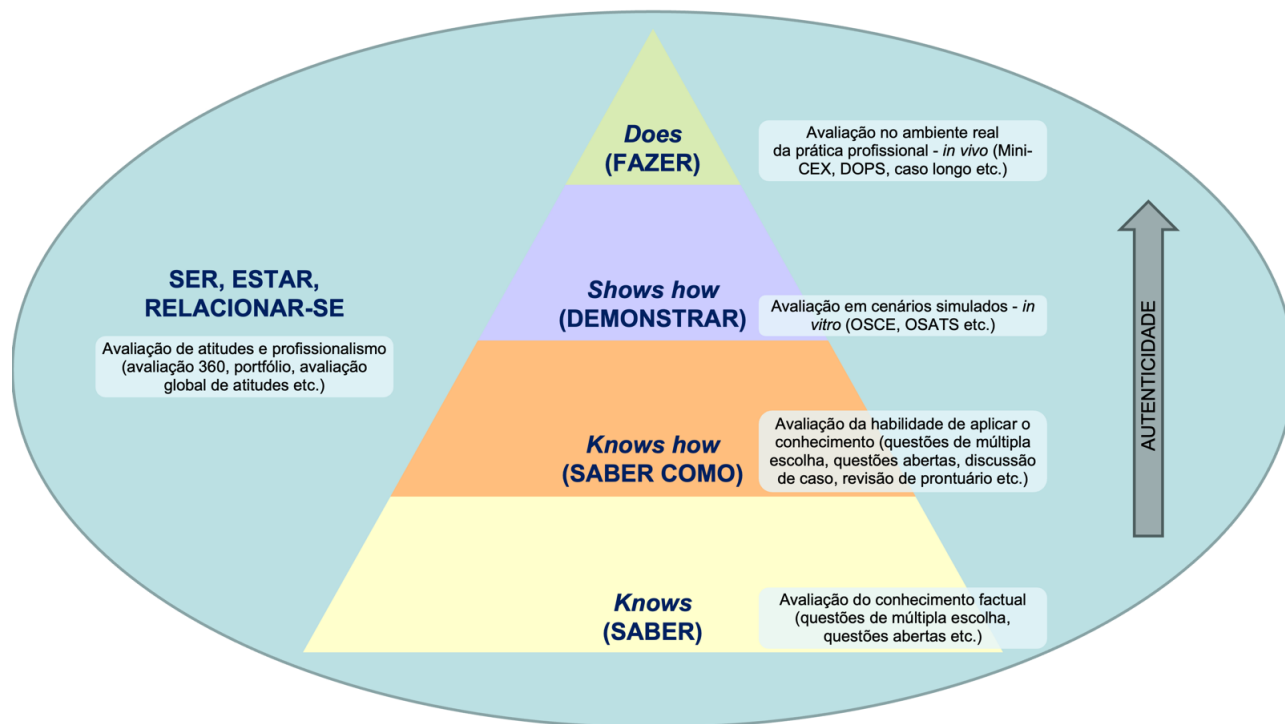
A avaliação do estudante: requisitos de qualidade

Independentemente do método utilizado, a avaliação do estudante precisa atender a atributos de qualidade já bem estabelecidos: validade, confiabilidade, viabilidade, aceitabilidade, impacto educacional e efeito catalisador^{2-4,7}. A validade diz respeito à adequação entre o que se pretende avaliar e o que efetivamente é observado. A confiabilidade refere-se à consistência das medidas obtidas. A viabilidade depende da existência de recursos humanos, materiais e organizacionais suficientes. A aceitabilidade envolve o grau de adesão de estudantes, docentes e instituições ao método empregado. Já o impacto educacional e o efeito catalisador correspondem à capacidade da avaliação de influenciar positivamente a aprendizagem, o ensino e o próprio currículo^{2,3,7}.

Foge ao escopo deste ensaio discutir todos os atributos, mas existem duas estratégias muito úteis na prática para garantir qualidade na avaliação: o uso da pirâmide de Miller (Figura 1) e o *blueprinting*. Em 1990, George Miller propôs um esquema para classificar a avaliação em quatro níveis: *knows* (saber), *knows how* (saber como), *shows how* (demonstrar como faz) e *does* (fazer)¹¹. A utilização desse esquema ajuda a alinhar o método de avaliação com o domínio a ser avaliado, conferindo validade. Nesse esquema, os níveis *knows* e *knows how* são mais adequadamente explorados por instrumentos cognitivos, ao passo que o nível *shows how* requer observação estruturada do desempenho em situações simuladas ou controladas.

A outra estratégia, o *blueprinting*, deriva da engenharia e consiste na construção de um mapa ou representação estruturada (*blueprint*) do que vai ser avaliado. Com o uso do *blueprinting*, é possível explicitar, já no planejamento da prova, a amostra de conteúdos, contextos e competências que serão avaliados. No OSCE, que avalia habilidades clínicas (e não especialidades), o *blueprinting* precisa assegurar representação equilibrada de competências clínicas essenciais, como anamnese, exame físico, comunicação, raciocínio clínico, tomada de decisão e profissionalismo¹²⁻¹⁴. Por exemplo, é mais relevante garantir uma distribuição adequada de estações que avaliem competências essenciais (como habilidade para realizar o exame físico, a comunicação e o raciocínio clínico) do que simplesmente organizar as estações em torno de áreas ou especialidades específicas (Figura 2).

Figura 1. Esquema representativo da pirâmide de Miller modificada. A modificação consiste em posicionar a pirâmide em um fundo circular que representa o domínio afetivo-atitudinal, expresso pelos verbos “ser”, “estar” e “relacionar-se”. Para cada nível da pirâmide e desse fundo, apresentam-se exemplos de métodos que podem ser utilizados, destacando-se o OSCE no nível do “demonstrar como faz”.



Fonte: Miller^{11,48}.

Figura 2. Exemplo de *blueprint* ou matriz de especificação para um OSCE destinado à avaliação de habilidades clínicas de formandos de um curso de graduação em Medicina. O exame propõe 14 estações, em cenários distintos, que avaliam seis grandes domínios de competências clínicas (e não especialidades). Nesse exame, cada competência é avaliada em contextos diversos, em um mínimo de cinco estações e um máximo de sete estações.

N	Área	Especialidade	Tema	Cenário	Comunicação	História clínica	Exame físico	Raciocínio clínico	Proposição e execução de ações (procedimentos)	Orientação do paciente e dos familiares	Observações/recursos necessários
1	Clínica médica	Cardiologia	Palpitações	Ambulatório especializado	X	X		X			Paciente simulado
2	Medicina de família	Não se aplica (geral)	Aferição de PA	Estratégia Saúde da Família	X				X	X	Paciente simulado
3	Clínica médica	Pneumologia	Ausculta de pneumonia	UPA	X		X	X			Paciente simulado e simulador
4	Ortopedia	Não se aplica (geral)	Dor lombar	UBS		X		X	X		Paciente simulado
5	Neurologia	Não se aplica (geral)	AVC	Pré-hospitalar		X	X	X			Paciente simulado
6	Pediatria	Puericultura	Vacinação	UBS	X	X				X	Paciente simulado e carteira de vacinação
7	Estação de descanso										
8	Ginecologia e obstetrícia	Não se aplica (geral)	Avaliação de trabalho de parto	Maternidade					X	X	Paciente simulado e simulador
9	Ginecologia e obstetrícia	Não se aplica (geral)	Método anticoncepcional	UBS	X			X		X	Paciente simulado
10	Pediatria	Endocrinologia	Diabetes <i>mellitus</i>	Ambulatório especializado	X	X		X			Paciente simulado
11	Medicina de família	Não se aplica (geral)	Mordedura de cão	Estratégia Saúde da Família		X	X			X	Paciente simulado e maquiagem
12	Otorrino	Não se aplica (geral)	Corpo estranho nasal	UPA			X		X		Paciente simulado e simulador
13	Cirurgia	Não se aplica (geral)	Cuidados pós-operatório	Hospital secundário	X		X			X	Paciente simulado e maquiagem
14	Cirurgia	Não se aplica (geral)	Sutura	Sala de procedimentos			X		X	X	Paciente simulado e modelo para sutura
Número de domínios avaliados					7	6	6	6	5	7	

Fonte: Elaborada pelos autores.

A combinação entre alinhamento conceitual e *blueprinting* adequado não elimina todos os problemas metodológicos, mas reduz ameaças importantes à validade e à equidade do processo. Em avaliações de maior consequência (*high-stakes exams*), esse planejamento precisa ser acompanhado por treinamento de avaliadores, definição explícita de critérios de aprovação, análise do desempenho das estações e monitoramento sistemático da variabilidade entre examinadores e circuitos^{7,12-14}. Em termos psicométricos, isso implica reconhecer que a confiabilidade do OSCE decorre menos do formato em si e mais da qualidade da sua implementação: número de estações, amostragem do domínio, padronização dos casos, comportamento dos avaliadores e estabilidade dos critérios de julgamento. Por isso, em exames de maior consequência, é desejável complementar indicadores tradicionais com abordagens mais robustas, como estudos de generalização, análise de consistência interavaliadores e revisão sistemática de estações com funcionamento insatisfatório.

A avaliação do estudante no contexto da formação baseada em competências

A educação médica contemporânea deslocou seu eixo de um currículo predominantemente centrado em conteúdo para uma formação baseada em competências⁵. Esse movimento alterou profundamente a lógica da avaliação. Já não basta verificar a retenção de informações ou a capacidade de resolver problemas em provas escritas; tornou-se necessário observar como o estudante integra conhecimentos, habilidades, comunicação, profissionalismo e tomada de decisão clínica em situações diretamente observáveis^{1,5-7}.

É precisamente nesse ponto que se localiza a maior contribuição do OSCE. Desde sua proposição, na década de 1970, o método permitiu estruturar a avaliação do desempenho clínico em estações com tarefas delimitadas, critérios explícitos e condições relativamente padronizadas. Com isso, reduziram-se algumas das limitações clássicas do exame à beira do leito, fortemente dependentes da complexidade de cada caso clínico, do observador e do contexto de aplicação⁸⁻¹⁰. Na pirâmide de Miller, o OSCE se situa predominantemente no nível *shows how*: ele permite observar o estudante colhendo história, realizando exame físico focalizado, explicando procedimentos, comunicando riscos e condutas, aconselhando pacientes e interpretando achados em condições comparáveis entre candidatos⁸⁻¹⁰.

Essa localização conceitual esclarece um ponto decisivo: o OSCE não avalia toda a competência clínica, mas uma amostra estruturada e intencional de desempenhos observáveis. Seu valor deriva exatamente dessa capacidade de amostrar, com razoável objetividade, facetas do construto competência

clínica que provas cognitivas não captam adequadamente. Em contrapartida, sua qualidade depende de escolhas rigorosas quanto ao propósito do exame, ao conteúdo das estações, ao treinamento de pacientes padronizados e avaliadores e ao uso apropriado de listas de verificação e escalas globais¹²⁻¹⁴. Sob essa ótica, o OSCE não deve ser interpretado como alternativa à avaliação no ambiente real de prática, mas como componente complementar de um sistema mais amplo. Avaliações em serviço documentadas, portfólios, supervisão longitudinal e julgamentos globais continuam indispensáveis para o nível *does*. O ganho específico do OSCE está em tornar comparável aquilo que, no cotidiano clínico, costuma ser exposto de forma desigual e oportunística para os estudantes.

Por essa razão, não existe “o” OSCE em abstrato. Há desenhos diferentes, apropriados a finalidades distintas. Em contextos formativos ou de baixa consequência (*low-stakes exams*), é possível trabalhar com circuitos menores, maior ênfase em *feedback* e menor exigência de confiabilidade global. Em contextos somativos ou de alta consequência (*high-stakes exams*), como progressão, certificação ou seleção, tornam-se indispensáveis maior amostragem de cada domínio, maior rigor psicométrico, processos formais de definição de nota de corte, controle de segurança e documentação detalhada do exame e de suas análises pós-implementação^{7,12-14}.

Introdução e difusão do OSCE no Brasil e na América Latina

A história internacional do OSCE é relativamente bem conhecida. Em Dundee, Ronald Harden e colaboradores formalizaram, em meados da década de 1970, um exame objetivo e estruturado para contornar limitações do exame clínico tradicional, mais tarde nomeado OSCE^{8,20,21}. Nas décadas seguintes, essa experiência consolidou-se como referência e foi ampliada para diferentes contextos, incluindo avaliação de habilidades técnicas, avaliação clínica de estudantes e residentes e processos de licenciamento médico na América do Norte²²⁻²⁷. Essa trajetória mostrou que o método podia ser integrado a currículos orientados por resultados e a sistemas de avaliação longitudinalmente distribuídos ao longo do curso^{10,28}.

Na América Latina, entretanto, a difusão do OSCE foi mais tardia e heterogênea. Uma revisão recente identificou aumento do uso do método, ainda concentrado em poucos países e frequentemente acompanhado de descrição metodológica incompleta, o que dificulta julgar a qualidade dos exames e comparar experiências²⁹. Essa fragilidade de relato não é trivial: sem informação suficiente sobre *blueprinting*, treinamento de avaliadores, características das estações, definição de padrões e análises de confiabilidade, a

aparente objetividade do OSCE pode induzir interpretações excessivamente otimistas sobre sua qualidade^{13,29}.

No caso latino-americano, a experiência brasileira merece destaque histórico. Entre 1989 e 1995, o Instituto de Pós-Graduação Médica Carlos Chagas, no Rio de Janeiro, participou, em parceria com a *Educational Commission for Foreign Medical Graduates* (ECFMG), de aplicações internacionais de exame clínico estruturado baseado em pacientes padronizados, com estações comuns em diferentes países²⁶. Embora essa experiência tenha sido pouco documentada posteriormente, ela indica que o Brasil participou cedo do movimento internacional de avaliação prática padronizada.

A experiência pioneira mais bem documentada ocorreu na Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (FMRP-USP). Em 1995, grupo liderado por Luiz Ernesto de Almeida Troncon implementou um OSCE de seis estações para avaliação final dos estudantes da disciplina de Semiologia Geral. Em 1996, a experiência foi publicada em dois artigos na Revista Brasileira de Educação Médica: o primeiro descreveu a construção do exame, com duas estações de anamnese em pacientes simulados e quatro de exame físico em pacientes reais e padronizados; o segundo apresentou os resultados da aplicação e a percepção dos 41 estudantes avaliados^{30,31}. Os autores assinalaram tratar-se, ao seu conhecimento, da primeira aplicação dessa modalidade no país.

Ainda na FMRP-USP, o OSCE passou a integrar a avaliação institucional terminal de graduandos em Medicina, inclusive em estudos de comparação curricular^{32,33}. Em paralelo, escolas com currículos inovadores, como a Universidade Estadual de Londrina e a Faculdade de Medicina de Marília, introduziram avaliações estruturadas de habilidades clínicas inspiradas no modelo OSCE, com adaptações locais coerentes com seus projetos pedagógicos^{34,35}. Já a difusão mais ampla do formato para processos seletivos ocorreu sobretudo após a inclusão de provas práticas nos concursos de residência médica, prevista pela Comissão Nacional de Residência Médica em 2004³⁶. Vale notar que essa trajetória brasileira não foi linear. Houve experiências institucionais de grande qualidade, mas também descontinuidade, heterogeneidade de nomenclatura e pouca publicação de dados de validade e confiabilidade, o que ajuda a explicar por que parte da produção nacional permaneceu pouco visível em revisões internacionais.

Lições aprendidas sobre o OSCE no Brasil e exemplos internacionais

As contribuições brasileiras iniciais foram relevantes não apenas por antecederem boa parte da literatura regional posterior, mas também porque tornaram visíveis problemas que permanecem atuais. Os estudos da FMRP-USP mostraram

que o OSCE produz informações detalhadas sobre o desempenho discente e pode revelar fragilidades curriculares com precisão superior à de avaliações cognitivas tradicionais. Ao mesmo tempo, expuseram desafios operacionais, culturais e institucionais importantes. Os estudantes avaliaram positivamente a pertinência dos casos e conteúdos, mas relataram dificuldades significativas com o controle do tempo e com o estresse inerente ao formato; já os docentes reconheceram a capacidade do método para avaliar habilidades específicas, mas destacaram seu custo elevado e limitações para captar uma abordagem mais integrada do paciente^{30,31,37}.

O estudo de 2004, também conduzido por Troncon³⁷, é particularmente instrutivo porque analisou três anos consecutivos de uso do OSCE em uma escola médica tradicional, envolvendo 258 estudantes e docentes da instituição. Mesmo após ajustes organizacionais, persistiram dificuldades relacionadas ao tempo por estação, ao estresse discente, ao consumo de recursos e à baixa incorporação institucional dos resultados. A principal lição talvez seja que a utilidade do OSCE não depende apenas de suas características intrínsecas, mas também do contexto em que é implementado. Um OSCE bem-sucedido exige governança educacional, apoio administrativo, capacitação docente e uma cultura institucional de utilização consequente dos resultados.

Essa conclusão ajuda a compreender a distinção entre uso formativo e somativo do método. Em contextos formativos, o OSCE pode ter grande impacto educacional quando articulado a *feedback* estruturado e a oportunidades reais de remediação. Nessa condição, exames menores e mais focalizados podem ser suficientes para orientar aprendizagem. Em contextos somativos e, sobretudo, de alta consequência, a exigência muda de patamar: é necessário ampliar a amostragem do domínio, reduzir fontes de variabilidade irrelevante, definir padrões de aprovação defensáveis e realizar análises pós-implementação mais robustas, inclusive com atenção ao comportamento de estações, avaliadores, *checklists* e escalas globais^{7,12-14}.

A questão torna-se ainda mais sensível quando se consideram exames de proficiência e licenciamento. No Brasil, a segunda etapa do Exame Nacional de Revalidação de Diplomas Médicos (Revalida) mantém a exigência de demonstração prática de competências por meio de estações estruturadas, reafirmando que o domínio cognitivo isolado é insuficiente para decisões regulatórias de alta relevância^{38,39}. Isso confere legitimidade à defesa de avaliações práticas também para egressos de escolas nacionais. Por sua vez, a experiência internacional recomenda cautela diante da ideia de um grande OSCE nacional centralizado. Nos Estados Unidos e no Canadá, componentes clínicos nacionais foram descontinuados, em contexto marcado por alto custo,

complexidade logística e fortalecimento gradual de avaliações clínicas realizadas ao longo da formação^{40,41}. Essas mudanças não significam perda de importância da avaliação prática. Ao contrário, sugerem que sistemas maduros tendem a distribuí-la de maneira mais inteligente.

Em contraste, outros países com tradição em avaliação e processos regulatórios bem amadurecidos mantiveram avaliações no formato OSCE, como Reino Unido, Austrália, Suíça e Japão, ainda que com desenhos distintos. Alguns países, como Suíça e Japão, mantêm exames com forte padronização nacional, embora em momentos diferentes da formação do médico; já outros, como Reino Unido e Austrália, utilizam OSCE de forma descentralizada, frequentemente com diretrizes e governança nacionais, mas sem uma aplicação única e centralizada para todos os candidatos.

O modelo do Reino Unido combina dois componentes: o *Applied Knowledge Test* e o *Clinical and Professional Skills Assessment*. Neste último, as próprias escolas médicas organizam e aplicam a avaliação de habilidades clínicas e profissionais, enquanto o regulador nacional exerce papel de garantia externa de qualidade, consistência e justiça do processo⁴²⁻⁴⁴. Esse arranjo demonstra que é possível articular referencial nacional comum com execução descentralizada, sem necessidade de reunir todos os concluintes do país em um único OSCE nacional.

Assim, em nossa opinião, o modelo britânico é particularmente relevante para o debate brasileiro, não como modelo a ser copiado mecanicamente, mas como evidência de que referencial nacional comum e execução distribuída podem coexistir de forma rigorosa. Em contrapartida, sua viabilidade depende de pré-requisitos institucionais robustos, entre eles acreditação, garantia externa de qualidade e monitoramento sistemático dos processos formativos. No Brasil, a expansão do número de cursos de Medicina sem um sistema obrigatório e robusto de acreditação torna esse desafio ainda maior^{45,46}. A própria literatura internacional mostra que a consolidação do OSCE de alta qualidade depende de ciclos continuados de desenho, teste, aplicação, revisão e redesenho; em outras palavras, não se trata apenas de aplicar uma prova prática, mas também de manter um programa de avaliação capaz de aprender com seus próprios dados^{7,14,42-44}.

Uma proposta para o Brasil

A criação do Enamed recolocou a discussão sobre como articular avaliação externa e certificação da formação médica no país^{16,18}. Embora o exame tenha, até o momento, foco predominantemente cognitivo, ele explicita a necessidade de um arranjo nacional coerente para a avaliação das competências clínicas ao final da graduação. A nosso ver, o principal equívoco

a evitar é a falsa oposição entre duas alternativas igualmente insatisfatórias: de um lado, um grande OSCE nacional único e centralizado para todos os concluintes; de outro, a ausência de um componente prático nacionalmente regulado.

Dadas as dimensões do sistema formador brasileiro, o número de escolas médicas, a heterogeneidade de infraestrutura e o volume anual de concluintes, um OSCE nacional centralizado parece logisticamente pouco viável e provavelmente excessivamente oneroso. Entretanto, essa inviabilidade não justifica renunciar à avaliação prática. Uma alternativa mais consistente pode ser um modelo híbrido: prova cognitiva nacional centralizada e avaliação das habilidades clínicas por meio de OSCE locais ou regionais, conduzidos por escolas médicas individualmente ou em consórcios, mas submetidos a uma matriz nacional de competências, a requisitos mínimos de *blueprinting*, a padrões documentais comuns e a mecanismos independentes de auditoria externa.

Esse modelo apresenta vantagens claras. Preserva a avaliação prática como requisito de certificação, favorece maior integração entre currículo, ensino e avaliação longitudinal, reduz a complexidade de um exame único de escala nacional e estimula cooperação interinstitucional para compartilhamento de estações, formação de avaliadores e treinamento de pacientes padronizados. Além disso, permite que parte do trabalho técnico seja centralizada ou compartilhada, como o desenvolvimento de um banco nacional de estações, guias de *blueprinting*, referenciais para definição de nota de corte e protocolos de análise pós-implementação. Também é importante reconhecer que a descentralização só será legítima se vier acompanhada de transparência e comparabilidade. Não basta que cada escola declare realizar um OSCE; será necessário demonstrar, com evidências auditáveis, quais competências foram avaliadas, em que condições, por quem, com quais instrumentos, com qual padrão de aprovação e com quais resultados agregados.

Para que esse arranjo seja crível, alguns elementos são indispensáveis (Quadro 1): uma matriz nacional de competências clínicas; documentação padronizada do planejamento e da execução dos exames; programas permanentes de desenvolvimento docente e de treinamento de pacientes padronizados; análise pós-implementação, incluindo definição de nota de corte, medidas de confiabilidade e estudo do comportamento das estações; e mecanismos externos de auditoria, regulação e, idealmente, acreditação^{7,14,43,45-47}.

Por fim, considerando o esforço hercúleo de estruturar e aprimorar a avaliação de habilidades clínicas em um país continental com mais de 450 escolas médicas, esse desafio não deve ser delegado a uma única instituição. Ele exige coordenação entre o Ministério da Educação, a Associação Brasileira de

Quadro 1. Elementos indispensáveis para conferir qualidade à proposta de implementação da avaliação de habilidades clínicas pelas escolas médicas inseridas em um contexto nacional de verificação da adequação da formação do graduando.

ETAPAS	ELEMENTOS
1	Criação de uma matriz nacional de habilidades e competências clínicas e proposição de um <i>blueprinting</i> mínimo para a sua avaliação.
2	Proposição de condições mínimas padronizadas que permitam às escolas documentar as principais etapas do planejamento e da execução dos exames.
3	Criação de programas locais de desenvolvimento docente e de treinamento de pacientes padronizados, com linha de trabalho específica para a avaliação de habilidades e competências clínicas.
4	Incorporação de uma sistemática de análise pós-execução dos exames (definição da nota de corte, medidas de confiabilidade, variabilidade interavaliadores e intra-avaliadores, medidas do impacto educacional etc.).
5	Constituição de instâncias externas de âmbito nacional para acompanhamento e monitoramento das ações (auditoria e regulação).

Fonte: Elaborado pelos autores.

Educação Médica, o Conselho Federal de Medicina, os órgãos reguladores, as escolas médicas, os estudantes e as entidades de acreditação. No entanto, considerando a complexidade de um exame como esse, é fundamental que exista também um grupo técnico independente, responsável por definir padrões de qualidade, acompanhar a implementação e revisar continuamente a qualidade do sistema.

CONCLUSÕES

A trajetória brasileira do OSCE oferece uma lição dupla. De um lado, demonstra que o país participou precocemente da introdução do método na América Latina e acumulou experiência suficiente para contribuir de forma qualificada para o desenho de soluções contemporâneas. De outro lado, evidencia que a simples adoção do formato não resolve, por si só, os problemas da avaliação clínica. Sem ambiente institucional favorável, investimento, rigor técnico e cultura de uso consequente dos resultados, o exame perde legitimidade e sustentabilidade. Em síntese, o OSCE permanece um instrumento central para a avaliação das competências clínicas na educação médica. Sua relevância atual, contudo, exige respostas proporcionais à escala e à diversidade do sistema formador brasileiro. Um arranjo híbrido, com avaliação cognitiva centralizada e avaliação prática descentralizada sob forte regulação nacional, parece ser a alternativa mais realista para conciliar validade, factibilidade e responsabilidade social na certificação do futuro médico.

CONTRIBUIÇÃO DOS AUTORES

Marcos Carvalho Borges, Valdes Roberto Bollela e Luiz Ernesto de Almeida Troncon contribuíram com escrita, revisão crítica e aprovação final da versão.

CONFLITO DE INTERESSES

Declaramos não haver conflito de interesses.

FINANCIAMENTO

Declaramos não haver financiamento.

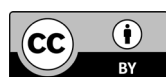
DECLARAÇÃO DE DISPONIBILIDADE DE DADOS

Os dados de pesquisa estão disponíveis no corpo do documento.

REFERÊNCIAS

- van der Vleuten CP, Schuwirth LW, Driessen EW, Dijkstra J, Tigelaar D, Baartman LK, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012;34(3):205-14.
- Panúncio-Pinto MP, Troncon LEA. Avaliação do estudante – aspectos gerais. *Medicina (Ribeirão Preto)*. 2014;47(3):314-23.
- Norcini J, Anderson B, Bollela V, Burch V, Costa MJ, Duvivier R, et al. Criteria for good assessment: consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach*. 2011;33(3):206-14.
- Troncon LEA. Avaliação do estudante de medicina. *Medicina (Ribeirão Preto)*. 1996;29:429-39.
- Frank JR, Snell LS, Cate OT, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education: theory to practice. *Med Teach*. 2010;32(8):638-45.
- Norcini J, Anderson MB, Bollela V, Burch V, Costa MJ, Duvivier R, et al. 2018 Consensus framework for good assessment. *Med Teach*. 2018;40(11):1102-9.
- Boursicot K, Kemp S, Wilkinson T, Findyartini A, Canning C, Cilliers F, et al. Performance assessment: consensus statement and recommendations from the 2020 Ottawa Conference. *Med Teach*. 2021;43(1):58-67.
- Harden RM, Stevenson M, Downie WW, Wilson GM. Assessment of clinical competence using objective structured examination. *Br Med J*. 1975;1(5955):447-51.
- Harden RM. What is an OSCE? *Med Teach*. 1988;10(1):19-22.
- Harden RM. Revisiting "Assessment of clinical competence using an objective structured clinical examination (OSCE)". *Med Educ*. 2016;50(4):376-9.
- Miller GE. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65(9 Suppl):S63-7.
- Boursicot K, Roberts T. How to set up an OSCE. *Clin Teach*. 2005;2(1):16-20.

13. Patricio M, Juliao M, Fareleira F, Young M, Norman G, Vaz Carneiro A. A comprehensive checklist for reporting the use of OSCEs. *Med Teach*. 2009;31(2):112-24.
14. Daniels VJ, Pugh D. Twelve tips for developing an OSCE that measures what you want. *Med Teach*. 2018;40(12):1208-13.
15. Oliveira SS, Rodrigues ET, Afonso DH, Ribeiro ACA, Palhares AA, Moraes NP, et al. Novas Diretrizes Curriculares Nacionais do curso de graduação em Medicina de 2025: um convite para a transformação da educação médica no Brasil. *Rev Bras Educ Med*. 2026;50(01):e050.
16. Ministério da Educação. Portaria MEC Nº 330, de 23 de abril de 2025. Institui o Exame Nacional de Avaliação da Formação Médica – Enamed. Brasília: MEC; 2025.
17. Flauzino RH, Panúncio-Pinto MP, Bollela VR, Troncon LEA. Avaliação do estudante de graduação nas profissões da saúde: mudanças associadas ao desenvolvimento docente. *Rev Bras Educ Med*. 2025;49(3):e089.
18. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Enamed. Brasília: Ministério da Educação; 2025 [acesso em 10 de maio de 2026]. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enamed>.
19. Ministério da Educação. Resolução CNE/CES nº 3, de 30 de setembro de 2025. Institui Diretrizes Curriculares Nacionais do Curso de Graduação em Medicina. Brasília: Ministério da Educação; 2025.
20. Harden RM. Assess clinical competence-an overview. *Med Teach*. 1979;1(6):289-96.
21. Harden RM, Gleeson FA. Assessment of clinical competence using an objective structured clinical examination (OSCE). *Med Educ*. 1979;13(1):41-54.
22. Faulkner H, Regehr G, Martin J, Reznick R. Validation of an objective structured assessment of technical skill for surgical residents. *Acad Med*. 1996;71(12):1363-5.
23. Stillman PL, Regan MB, Swanson DB, Case S, McCahan J, Feinblatt J, et al. An assessment of the clinical skills of fourth-year students at four New England medical schools. *Acad Med*. 1990;65(5):320-6.
24. Stillman P, Swanson D, Regan MB, Philbin MM, Nelson V, Ebert T, et al. Assessment of clinical skills of residents utilizing standardized patients. A follow-up study and recommendations for application. *Ann Intern Med*. 1991;114(5):393-401.
25. Papadakis MA. The Step 2 clinical-skills examination. *N Engl J Med*. 2004;350(17):1703-5.
26. Ziv A, Ben-David MF, Sutnick AI, Gary NE. Lessons learned from six years of international administrations of the ECFMG's SP-based clinical skills assessment. *Acad Med*. 1998;73(1):84-91.
27. Reznick RK, Blackmore D, Cohen R, Baumber J, Rothman A, Smee S, et al. An objective structured clinical examination for the licentiate of the Medical Council of Canada: from research to reality. *Acad Med*. 1993;68(10 Suppl):S4-6.
28. Harden RM. AMEE Guide No. 14: Outcome-based education: Part 1-An introduction to outcome-based education. *Med Teach*. 1999;21(1):7-14.
29. Armijo-Rivera S, Fuenzalida-Munoz B, Vicencio-Clarke S, Elbers-Arce A, Bozzo-Navarrete S, Kunakov N, et al. Advancing the assessment of clinical competence in Latin America: a scoping review of OSCE implementation and challenges in resource-limited settings. *BMC Med Educ*. 2025;25(1):587.
30. Troncon LEA, Foss NT, Voltarelli JC, Dantas RO. Avaliação de habilidades clínicas por exame objetivo estruturado por estações, com emprego de pacientes padronizados: descrição de dois métodos (parte I). *Rev Bras Educ Med*. 1996;20(2-3):45-52.
31. Troncon LEA, Foss NT, Voltarelli JC, Dantas RO. Avaliação de habilidades clínicas por exame objetivo estruturado por estações, com emprego de pacientes padronizados; uma aplicação no Brasil (parte II). *Rev Bras Educ Med*. 1996;20(2-3):53-60.
32. Figueiredo JF, Troncon LE, Rodrigues M de L, Cianflone AR, Colares M de F, Peres LC, et al. Effect of curriculum reform on graduating student performance. *Med Teach*. 2004;26(3):244-9.
33. Troncon LEA, Figueiredo JFC, Rodrigues MLV, Piccinato CE, Peres LC, Cianflone ARL, et al. Avaliação de uma reestruturação curricular na Faculdade de Medicina de Ribeirão Preto: influência sobre o desempenho dos graduandos. *Rev Bras Educ Med*. 2004;28(02):145-55.
34. Perim G, Sakai M, Almeida M, Marchese M. Sistema Integrado de Avaliação do Curso de Medicina da UEL: SIAMed. Avaliação: Revista da Avaliação da Educação Superior (Campinas). 2005;10:135-69.
35. Mazzoni C, Moraes M. A avaliação prática estruturada de habilidades clínicas na Famema: fundamentos para construção e aplicação. *Gestão Univ*. 2006;99:63-70.
36. Comissão Nacional de Residência Médica. Resolução CNRM nº 008/2004, de 05 de agosto de 2004. Dispõe sobre o processo de seleção pública dos candidatos aos Programas de Residência Médica. Brasília: CNRM; 2024.
37. Troncon LEA. Clinical skills assessment: limitations to the introduction of an "OSCE" (Objective Structured Clinical Examination) in a traditional Brazilian medical school. *São Paulo Medical Journal*. 2004;122(1):12-7.
38. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Revalida. Brasília: Ministério da Educação; 2026 [acesso em 10 de maio de 2026]. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/revalida>.
39. Ministério da Saúde, Ministério da Educação. Matriz de correspondência curricular para fins de revalidação de diplomas de médico obtidos no exterior. Brasília: MEC, MS; 2009.
40. Howley LD, Engle DL. Discontinuation of the USMLE Step 2 Clinical Skills Examination: studying the past to define the future. *Acad Med*. 2021;96(9):1247-9.
41. Smirnova A. Licensing exams in Canada: a closer look at the validity of the MCCQE Part II. *Can Med Educ J*. 2022;13(4):23-9.
42. General Medical Council. Clinical and professional skills assessment (CPSA). London: General Medical Council; 2026 [acesso em 10 de maio de 2026]. Disponível em: <https://www.gmc-uk.org/education/medical-licensing-assessment/uk-medical-schools-guide-to-the-mla/clinical-and-professional-skills-assessment-cpsa>.
43. General Medical Council. Requirements for the Medical Licensing Assessment Clinical and Professional Skills Assessment. London: General Medical Council; 2024.
44. General Medical Council. Promoting excellence: standards for medical education and training. London: General Medical Council; 2026.
45. Bedoll D, van Zanten M, McKinley D. Global trends in medical education accreditation. *Hum Resour Health*. 2021;19(1):70.
46. Tempiski P, Girotto LC, Brenelli S, Giamberardino DD, Martins MA. Accreditation of medical education in Brazil: an evaluation of seventy-six medical schools. *BMC Med Educ*. 2024;24(1):656.
47. Australian Medical Council. Procedures for Assessment and Accreditation of Medical Schools by the Australian Medical Council. Kingston: Australian Medical Council; 2024.
48. Bollela VR, Borges MC, Volpe FAP, Santos LL, Santana RC, Ricz HMA, et al. Princípios gerais de avaliação do profissional da saúde em formação. Residência médica : ensino e avaliação das competências: Manole; 2022.



This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.